DT/KNN2 (Q10L11) 25% (5/20)

✗ 1. This function can be used to perform KNN classificationin R

- A knn()
- B k_nn()
- C knnreg()
- D knearneib()
- E I do not know

✗ 2. With the increase of k, the decision boundary will be

- A simplified
- B more complex
- C I do not know
- D unchanged

✓ 3. In the case of small k we have

- A overfitting
- B underfitting
- C it depends on the situation
- D I do not know

✓ 4. Do you need to worry about scaling with one explanatory variable?

- A No
- B Yes
- C I do not know

✗ 5. n - the number of observation,
m - the number of explanatory variables

When n=k, m=1, the decision boundary for regression is

- A a line
- B a stepwise constant function
- C a stepwise quadratic function
- D I do not know

✗ 6. Which of these algorithms can be used to fill the missing values

- (A) KNN for regression
- (B) KNN for classification
- (C) both
- (D) I do not know

✗ 7. Which one is better: KNN regression or Linear regression ?

- (A) KNN outperform LR if the parametric form that has been selected is close to the true form of f
- (B) LR outperform KNN if the parametric form that has been selected is close to the true form of f
- (C) KNN will always outperform the LR
- (D) I do not know

✓ 8. Which one is the Disadvantage of KNN?

- (A) required assumptions
- (B) cannot be applied for regression
- (C) difficult to perform
- (D) the problem of high dimensional data
- (E) I do not know

✓ 9. The best k for train set equals to

- (A) 1
- (B) 2
- (C) 0
- (D) I do not know

✗ 10. Decision tree is

- (A) supervised learning algorithm
- (B) unsupervised learning algorithm
- (C) I do bot know

✗ 11. Decision Tree Decision Boundaries

- (A) are a step-wise constant function
- (B) I do not know
- (C) continuous function
- (D) are axis-parallel rectangles

✓ **12. Root Node has**

- A no incoming edges and zero or more outgoing edges
- B one incoming edge and two or more outgoing edges
- C one incoming edge and no outgoing edges
- D I do not know

✗ **13. Child or Internal Node has**

- A no incoming edges and zero or more outgoing edges
- B one incoming edge and two or more outgoing edges
- C one incoming edge and no outgoing edges
- D I do not know

✗ **14. Pruning the tree means**

- A Simplify the tree
- B Split the tree's nodes
- C Merge the tree's nodes
- D I do not know

✗ **15. Gini index equals to**

- A 1 - sum (pi^2)
- B 1 + sum (pi^2)
- C sum(pi * log(pi))
- D -sum(pi * log(pi))
- E I do not know

✗ **16. Entropy starts with 0**

- A True
- B False
- C I do not know

✗ **17. Overall impurity measure can be obtained by**

- A a weighted average of individual rectangles
- B majority voting
- C I do not know

✗ 18. At each stage, we choose the split with

- A the lowest Gini index
- B the lowest Chi-square value
- C the highest entropy
- D I do not know

✗ 19. We can perform the Decision Trees in r using

- A rpart()
- B decisiontree()
- C destree()
- D reg.tree()
- E I do not know

✗ 20. minsplit in R means

- A the minimum number of observations that must exist in a node in order for a split to be attempted
- B the minimum number of observations in any terminal node
- C the minimum number of splits
- D I do not know